



Estimating multiple greenspace exposure types and their associations with neighbourhood premature mortality: A socioecological study

S.M. Labib^{a,b,*}, Sarah Lindley^b, Jonny J. Huck^b

^a Centre for Diet and Activity Research, MRC Epidemiology Unit, University of Cambridge, United Kingdom of Great Britain and Northern Ireland

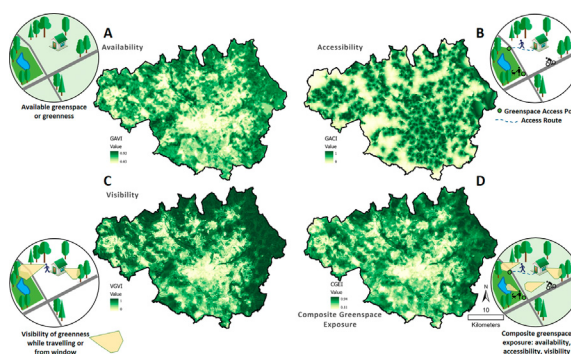
^b MCGIS, Department of Geography, School of Environment, Education and Development (SEED), University of Manchester, United Kingdom of Great Britain and Northern Ireland



HIGHLIGHTS

- Existing greenness exposure metrics are inadequate in estimating multiple pathways.
- We developed novel metrics of greenspace availability, accessibility, and visibility.
- We generated composite greenspace exposure combining three individual metrics.
- We found inverse associations between composite greenspace and premature mortality.
- The positive impact of greenspace exposure is more pronounced in deprived neighbourhoods.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:

Received 14 March 2021

Received in revised form 14 May 2021

Accepted 15 May 2021

Available online 21 May 2021

Editor: Jay Gan

Keywords:

Green space

Greenness

Pathways

Mortality

Exposome

Geographic information systems

ABSTRACT

Background: Greenspace exposures are often measured using single exposure metrics, which can lead to conflicting results. Existing methodologies are limited in their ability to estimate greenspace exposure comprehensively. We demonstrate new methods for estimating single and combined greenspace exposure metrics, representing multiple exposure types that combine impacts at various scales. We also investigate the association between those greenspace exposure types and premature mortality.

Methods: We used geospatial data and spatial analytics to model and map greenspace availability, accessibility and eye-level visibility exposure metrics. These were harmonised and standardised to create a novel composite greenspace exposure index (CGEI). Using these metrics, we investigated associations between greenspace exposures and years of potential life lost (YPLL) for 1673 neighbourhoods applying spatial autoregressive models. We also investigated the variations in these associations in conjunction with levels of socioeconomic deprivation based on the index of multiple deprivations.

Results: Our new CGEI metric provides the opportunity to estimate spatially explicit total greenspace exposure. We found that a 1-unit increase in neighbourhood CGEI was associated with approximately a 10-year reduction in YPLL. Meaning a 0.1 increment or 10% increase in the CGEI is associated with an approximately one year lower premature mortality value. A single 1-unit increase in greenspace availability was associated with a YPLL reduction of 9.8 years, whereas greenness visibility related to a reduction of 6.14 years. We found no significant association between greenspace accessibility and YPLL. Our results further identified divergent trends in the relations

Abbreviations: ANGSt, Accessible Natural Greenspace Standards; CGEI, Composite Greenspace Exposure Index; GACI, Greenspace Accessibility Index; GAVI, Greenspace Availability Index; GIS, Geographical Information Science; GLM, Generalised Linear Model; GM, Greater Manchester; IMD, Index of Multiple Deprivation; LSOA, Lower Super Output Area; LULC, Land Use and Land Cover; MAUP, Modifiable Areal Unit Problem; NDVI, Normalised Difference Vegetation Index; SARAR, Spatial Autoregressive With Autoregressive Error; UGCoP, Uncertain Geographic Context Problem; VGV, Viewshed Greenness Visibility Index; YPLL, Years Of Potential Life Lost.

* Corresponding author at: Centre for Diet and Activity Research, MRC Epidemiology Unit, University of Cambridge, United Kingdom of Great Britain and Northern Ireland.

E-mail addresses: sml80@medschl.cam.ac.uk, labib.l.m@gmail.com (S.M. Labib), sarah.lindley@manchester.ac.uk (S. Lindley), jonathan.huck@manchester.ac.uk (J.J. Huck).

between greenspace exposure types (e.g. availability vs. accessibility) and levels of socioeconomic deprivation (e.g. least vs. most).

Conclusion: Our methods and metrics provide a novel approach to the assessment of multiple greenspace exposure types, and can be linked to the broader exposome framework. Our results showed that a higher composite greenspace exposure is associated with lower premature mortality.

© 2021 The Author. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The connection between human health and the natural environment is widely recognised (Frumkin et al., 2017; Díaz et al., 2018; Bratman et al., 2019; Remme et al., 2021). Several recent reviews have demonstrated that human interactions with nature, particularly exposure to green and blue spaces (commonly referred to jointly as 'greenspaces') provide multiple pathways enabling positive health benefits (Hartig et al., 2014; Markevych et al., 2017). The importance of these pathways is also recognised in the increasing emphasis on urban green infrastructure as a core part of spatial planning (Lindley et al., 2019; Dennis et al., 2020). Existing literature indicates that exposure to greenspace reduces health risk factors associated with all-cause mortality, adverse pregnancy outcomes, elevated blood pressure, type 2 diabetes, among others via improved physical health (Kondo et al., 2018; Rojas-Rueda et al., 2019; Jimenez et al., 2021). Studies have also reported positive effects related to greenspace exposure on mental health. Positive effects on mental health include the reduction of anxiety levels and stress levels as well as alleviating depression, improving sleep, enhancing cognitive development, and improving overall life satisfaction (Dadvand et al., 2015; Fong et al., 2018; Kondo et al., 2018). Nonetheless, a few studies have also found null, mixed, or adverse effects related to greenspace exposure (Picavet et al., 2016; Dzhambov et al., 2020). Noting this, some scholars have argued that the possible reasons for these inconsistent findings might be attributed to differences in measuring exposures in terms of both spatial and temporal variations as well as in the context of specific health and medical conditions (e.g., allergies) that may be negatively impacted by some types of green space exposure (Markevych et al., 2017; Davis et al., 2021; Stas et al., 2021).

The conceptualisation of greenspace exposure used to frame previous studies involving objective measurement can be broadly classified into three categories; greenspace availability, accessibility, and visibility (Dadvand and Nieuwenhuijsen, 2019). These three categories represent the most commonly used spatial measures of greenspace exposure (Ekkel and de Vries, 2017; Labib et al., 2020b). Each of these three greenspace exposure categories are linked to different, though often overlapping, mechanistic pathways influencing health (Hartig et al., 2014; James et al., 2015; Markevych et al., 2017; Bratman et al., 2019). For example, availability of greenspace has been associated with the reduction of environmental stressors (e.g., air and noise pollution, heat mitigation) (Frumkin et al., 2017; Lindley et al., 2019). Access to greenspaces may encourage physical activities and increase social cohesion, linked to the capacity building pathway (Hartig et al., 2014; Markevych et al., 2017; Nieuwenhuijsen et al., 2017). Visibility of greenspace or 'greenness' visibility (Labib et al., 2021) may be associated with stress recovery and attention restoration (Ulrich, 1984; Kaplan and Kaplan, 1989). Additionally, spatial measures of greenspace exposure may act as proxies for the "cumulative opportunity" of exposure to nature (Frumkin et al., 2017), the actual exposure will depend on the usage and time spent in contact with greenspaces (Bratman et al., 2019; Holland et al., 2021). It should be noted that, unlike some other external environmental exposure types (Turner et al., 2017), the precise impact and value of greenspace exposure influences are, as yet, unclear. This is due at least in part to the multiple mechanistic pathways involved in health benefits, the heterogeneous exposure assessment methods applied in existing studies, and the difficulty in

accounting for actual exposure in terms of amount of time spent in nature (Turner et al., 2017; Labib et al., 2020b; Holland et al., 2021).

Existing approaches to the objective measurement of greenspace exposure typically only consider *one* of these three types of exposure (availability, accessibility, and visibility) and often associate them individually with health indicators. Accordingly, the aggregate value of the impact of greenspace on human health is not yet well understood (Frumkin et al., 2017; Bratman et al., 2019; Labib et al., 2020b). Understanding the influence of aggregate exposure is crucial because many of the non-communicable disease and health indicators considered in relation to greenspace exposure can be attributed to the aggregate effects of multiple exposure types (Frumkin et al., 2017; Silva et al., 2018; Labib et al., 2020b). Additionally, individuals are exposed to varying amounts and types of greenspace at different locations (e.g., home, office, streets), and this variability may result in differing health effects both on any given individual and from individual to individual over time (Helbich, 2018). Such variations in exposure over different spatial and temporal extents also relates to the uncertain geographic context problem (UGCoP) noted by (Kwan, 2012). For greenspace exposure, the UGCoP indicates that there is spatial uncertainty in the actual areas over which the greenspace exposure can impact health outcomes along multiple pathways and contexts (Pearce, 2018; Labib et al., 2020a). The pre-existing methodologies applied to greenspace exposure assessments are usually considered at a fixed spatial scale. As a result, there are uncertainties in estimating the aggregate effects of greenspace exposure at multiple locations related to varying temporal and spatial scales (Kwan, 2012; Helbich, 2018; Labib et al., 2020b). To better understand exposure duration, exposure sequences, and exposure accumulation, measurement of aggregate greenspace exposures at multiple scales therefore requires careful attention.

In this regard, new methodological solutions are required that allow the measurement of composite greenspace exposure based on hierarchical spatial scale values. Turner et al. (2017) indicated that conducting assessments of greenspace exposures that represent the totality of all such measurements in aggregate may have only moderate feasibility for the study of larger populations. Turner et al. also noted that it would be difficult to infer this exposure at the individual-level in such populations. However, the rapid advancement in spatial science disciplines (e.g., Geographical Information Science- GIS) and related location-based technologies, such as remote sensing, Global Navigation Satellite Systems, and environmental and personal sensors, in conjunction with increased computing capabilities for big data analysis could help overcome at least some of the aforementioned challenges (Jia, 2019; Labib et al., 2020b).

Considering the above arguments, the primary aim of the present study was to develop a new index that allows estimation of aggregate greenspace exposures at hierarchical spatial scales by applying tools and techniques related to GIS, satellite images, and other spatial data. We present a new composite greenspace exposure index (CGEI) that integrates three spatially explicit objective greenspace exposure types (i.e., availability, accessibility, and visibility) for the provision of city- or region-wide measurements of greenspace exposure at fine spatial resolutions. We also investigate statistical associations between single and composite greenspace exposure estimates at the neighbourhood scale in relation to years of potential life lost (YPLL) as a measure of premature mortality. Previous studies indicated that the availability of greenspace reduced mortality and morbidity for diverse populations

(Coutts and Horner, 2016; James et al., 2016; Fong et al., 2018; Rojas-Rueda et al., 2019), and that greenspace availability varies with the level of social deprivation and inequality in any given neighbourhood (Mitchell and Popham, 2008; Labib et al., 2020a). We therefore hypothesise that greenspace exposure characteristics (and by extension YPLL) for socially deprived neighbourhoods will differ depending on whether composite or single greenspace exposure measurements are used. We also expect that overall greenspace exposure in such neighbourhoods will reduce YPLL in those neighbourhoods.

2. Methods and materials

2.1. Study settings

We undertook this study in the Greater Manchester (GM) area of the United Kingdom. Greater Manchester is a post-industrial city-region with an area of 1276 km² and an estimated population of 2.8 million as of 2016 (Dennis et al., 2018). The city-region has a diverse landscape pattern, including flat plain areas surrounded by hills (the Pennine Chain) that rise to 500 m in the north and east. It has varying greenspaces of different sizes and types such as urban parks, pastoral areas, river corridors, and multiple areas of urban to rural transition (Dennis et al., 2018; Labib et al., 2020a). In the context of greenspace exposure, the differing characteristics in the area's natural environment are of particular interest for understanding the variability and importance of different types of greenspace exposure. From the perspective of both human health and socioeconomic inequalities, Greater Manchester is also an important area of study in a UK context, because neighbourhoods in the city-region have a range of social deprivation characteristics, from the particularly affluent through to extremely deprived. These levels, in turn, are associated with both varied and elevated levels of illness and disability (GMCA, 2019; Dennis et al., 2020).

2.2. Single greenspace exposure modelling

The composite greenspace exposure index we created in this study represents a combined measurement of three single greenspace exposure metrics (i.e., availability, accessibility, and visibility). We developed each of these individual metrics as distinct measures of greenspace exposure, with each following a different methodological approach using multiple sources of spatial data, tools, and analytical frameworks. The following sections summarise and discuss the methods applied in developing the single exposure metrics.

2.2.1. Greenspace availability exposure

We created a greenspace availability exposure index (GAVI) as a multi-scale, multi-metric map combining three commonly used greenspace metrics i.e., Normalised Difference Vegetation Index (NDVI), Leaf Area Index (LAI), and Land Use-Land Cover (LULC) at five spatial scales (i.e., 100, 200, 300, 400, and 500 m spatial resolutions - which can be thought of as equivalent to differing buffer distances commonly used to estimate exposure) following the methodology described in Labib et al. (2020a). These thresholds were carefully selected considering the spatial scale and resolution of the input data, and are thus less vulnerable to scaling effects (Labib et al., 2020a). We considered these three metrics to account for different characteristics of greenspace representing photosynthetically active vegetation. NDVI characterises greenspace density, LAI measures greenspace volume, and LULC accounts for the overall presence or absence of greenspace (James et al., 2015; Engemann et al., 2019). NDVI and LAI were obtained from 10 m resolution Sentinel-2 satellite images (collected on 4th July 2018, with cloud cover <2%). Sentinel-2 satellite images were selected due to their higher overall accuracy in identifying vegetation compared to moderate resolution Landsat images (Markevych et al., 2017; Labib and Harris, 2018). The LULC greenspace data were obtained from Dennis et al. (2018). Dennis et al. (2018) combined a supervised

classification of Sentinel 2A data from the Copernicus Scientific Data Hub with a range of supplementary data on land cover and land use, including from the Ordnance Survey Greenspace Layer. The LULC data included five land cover types (i.e., urban, water, tree canopy, forbs-shrubs, and grass), with a spatial resolution of 10 m.

For each greenspace metric, we calculated mean values in each grid cell of the map at five spatial scales (i.e., resolution) and combined these five scale maps using a weighted average method to produce multi-scale greenspace availability maps based upon NDVI, LAI and LULC (Fig. S1, *Supplementary document*). The multi-scale weighting, based upon lacunarity analysis, accounts for scale sensitivity effects observed at varying buffer distances, a detailed description of which is available in Labib et al. (2020a). Finally, we took the mean of the three maps to produce our multi-metric, multi-scale availability index using Eq. (1).

$$GAVI_j = \frac{(GS_{NDVIj} + GS_{LAIj} + GS_{LULCj})}{T} \quad (1)$$

In Eq. (1), $GAVI_j$ is the greenspace availability index value for cell j , and GS_{NDVIj} , GS_{LAIj} , and GS_{LULCj} are multi-scale greenspace metric 'exposure' values for corresponding cell j . T is the number of metrics (in our case, $T = 3$). In this case, $GAVI$ values ranged between 0 and 1, where 1 indicated the highest availability of greenspace, and 0 denoted the lowest level of, or no available, greenspace. It should be noted that we removed negative NDVI values from our computation, as these values represented bodies of water or other non-green features. We then set the negative NDVI values to zero. LAI values were normalised using minimum-maximum normalisation to obtain values consistent with other metrics. Details of LAI transformation and calculation can be found in Labib et al. (2020a). For this research, we re-sampled the GAVI map to 5 m resolution using nearest neighbour technique in ArcGIS Pro (v 2.4) to achieve consistency with other exposure layers.

2.2.2. Greenspace accessibility exposure

We computed the greenspace accessibility exposure index following a four-step method. First, we selected publicly accessible greenspaces within the study area. Second, we identified the access points for these greenspaces. Third, we ran a network analysis to measure and map the shortest network distance from the access points to any location within the study area. Fourth, we normalised the access distance to produce an accessibility raster map for the study area. The details of these steps follow.

2.2.2.1. Step 1: selecting accessible greenspaces. To determine accessible greenspaces, we used two GIS datasets; UK Ordnance Survey (OS) Open Greenspace (scale: 1:2500), and the OS Master Map Accessible Natural Greenspace layer (scale: 1:2500). We selected greenspaces that were publicly accessible to any users within or surrounding the GM area (up to 10 km). This distance was selected following Natural England's 'Accessible Natural Greenspace Standards' (ANGSt) model, which details standard access distances for varying greenspaces sizes in the UK see Comber et al. (2008). We considered the greenspaces outside the boundary of the study area to account for edge effects, whereby people living by the boundary may have access to nearby greenspaces that are just outside of, or on, the boundary. Our selection included public parks and open gardens, national parks, country parks, local nature reserves, and Countryside and Rights of Way Act (CRoW) accessible greenspaces. A detailed description of the greenspaces is listed in Table S1, *Supplementary document*. We selected these greenspace types based on previous studies evaluating greenspace accessibility in the UK (i.e., including spaces which are generally open to the public; (Pauleit et al., 2003; Comber et al., 2008). The size of our selected greenspaces ranged between 0.04 ha (400 m²) to more than 500 ha (5 km²), in order to ensure that we accounted for many forms of greenspace, from small 'pocket parks' through to large country parks.

2.2.2.2. Step 2: Access point identification. In urban areas like GM, greenspaces are usually entered through specific access points (Comber et al., 2008). To account for this, we identified the access points for greenspaces using two approaches. Firstly, using the OS Open Greenspace access points layer, we selected the access points that intersect the boundary of the greenspaces. However, this data layer did not contain access points for many natural greenspaces (e.g., country parks). To identify the missing access points, we performed a geometric intersection between the greenspace polygon layer (Step 1) and a road and path network layer (OS MasterMap ITN Layer Urban Paths, Scale 1: 1250) using ArcGIS Pro (v2.4). This intersection operation identified the points where streets and paths intersected the greenspace boundaries and, thus, could be access points. The two datasets were combined for use in this analysis.

2.2.2.3. Step 3: Network distance analysis. To determine the accessibility of a greenspace from a given location, we calculated the network distance, as this recognises that individuals generally need to follow designated road or path networks in order to reach greenspaces in urban areas (Labib et al., 2020b). This was achieved by undertaking a 'service area' analysis in QGIS (v 3.14) using the QNEAT3 toolbox to measure and map the network distance from greenspace access points to all other locations within the study area, using the QNEAT3 Isochrone Area (Iso-Area) algorithm. In Iso-Area, the starting points were represented by greenspace access points (identified in Step 2). Using the road and path network layer (OS ITN Layer Urban Paths; details: <https://bit.ly/31GaYlu>) the Iso-Area algorithm produced a network distance surface (raster) from the access points by applying the TIN-interpolation method (toolbox details: <https://root676.github.io/IsoAreaAlgs.html>) to a maximum distance of 10 km based on the upper limit in the ANGSt model (Pauleit et al., 2003; Comber et al., 2008). The output raster layer mapped the network distances at a spatial resolution of 10 m for computational efficiency. We masked the raster layer to exclude the extra cells generated outside the Greater Manchester boundary and re-sampled to 5 m resolution to achieve consistency with other layers and datasets.

2.2.2.4. Step 4: accessibility index calculation. Utilising the raster grid layer produced in Step 3 (Fig. S2, Supplementary document), we computed the greenspace accessibility index for each raster grid cell within the study area using Eq. (2).

$$GACI_j = \frac{(ND_{max} - ND_j)}{ND_{max}} \quad (2)$$

In Eq. (2), $GACI_j$ is the greenspace accessibility index expressed as a continuous value for cell j , ranging between 0 and 1, where 0 indicates the lowest accessibility and 1 indicates the highest. ND_{max} is the maximum network distance from the nearest access point within the entire study area (in our case the maximum distance was 5932.7 m); ND_j is the network distance of a cell from the nearest greenspace access points.

2.2.3. Greenness visibility exposure

We measured and mapped eye-level greenness visibility exposure for an observer located at ground level by applying viewshed analysis using new software we developed as part of our analysis (available at: <https://github.com/jonnyhuck/green-visibility-index>). We combined a binary greenness layer produced using LULC data from Dennis et al. (2018) with LiDAR based digital surface and digital terrain model data (Environmental Agency, 2016) to perform viewshed analysis at a 800 m viewing distance at 5 m intervals for >86 million observer locations covering the whole study area. We used a line of sight (LOS) algorithm, in conjunction with an empirical distance decay model, in order to calculate a distance-weighted ratio of visible greenness to visible area from each given observation point. The details of the decay model and LOS algorithm are reported in Labib et al. (2021). Based on

the LOS and decay weight values, we used Eq. (3) to compute the viewshed greenness visibility index for the study area. The details of the data, algorithms, steps, and modelling process can be found in Labib et al. (2021).

$$VGVI_j = \frac{\sum_1^n G_i \times d_i}{\left(\sum_1^n G_i \times d_i\right) + \left(\sum_1^n V_i \times d_i\right)} \quad (3)$$

where, $VGVI_j$ is the 'viewshed greenness visibility index' value for the observer cell j ; G_i is the visible green cell, V_i is the visible non-green cell, and d_i is distance decay weight corresponding to visible cell i . The estimated $VGVI$ values ranged between 0 and 1, where 0 = no visible green, and 1 = all visible green (Labib et al., 2021).

2.3. Composite greenspace exposure index (CGEI)

We combined the three single exposure indices into a composite index value for greenspace exposure. The resulting composite greenspace exposure index is produced by calculating the mean of the $GAVI$, $GACI$, and $VGVI$ datasets. We used Eq. (4) to produce the CGEI map.

$$CGEI_j = \frac{(GAVI_j \times w_{GAVI} + GACI_j \times w_{GACI} + VGVI_j \times w_{VGVI})}{(w_{GAVI} + w_{GACI} + w_{VGVI})} \quad (4)$$

In Eq. (4), $CGEI_j$ is the composite greenspace exposure index value for cell j ; $GAVI_j$, $GACI_j$, and $VGVI_j$ represent the respective availability, accessibility, and visibility index values for cell j obtained from single exposure maps; w represents the weighting value for each metric. The resulting operation produced the composite greenspace exposure index map at 5 m spatial resolution for the study area. Note that in this study we have used an equal weighting ($w = 1$) for each of the three single exposure metrics. Equal weighting was used in this case as currently there is no empirical evidence supporting alternative weightings for these single greenspace metrics in relation to their potential effects on premature mortality. However, the CGEI can also be estimated using differing weights for each individual metric. To illustrate the CGEI estimation process using different weightings for individual metrics we experimented with several hypothetical weight values see Supplementary note-1 for further details.

2.4. Assessment of CGEI and single exposure indices

The potential implications of the composite greenspace exposure metric reported in the present study were evaluated in two ways. Firstly, we investigated how the composite greenspace exposure measurement and its individual components each related to neighbourhood socioeconomic deprivation. Secondly, we compared each of the metrics to neighbourhood YPLL. We used English Lower Super Output Area (LSOA) Census units due to the availability of population-level socioeconomic deprivation and YPLL data, as well as several confounding variables. LSOAs contain a mean population of around 1500 and range from 1000 to 3000 and are widely used to represent neighbourhood units in national analyses (Mitchell and Popham, 2008; Daras et al., 2019). In our analysis, we examined 1673 LSOAs with a total population of 2,682,528 and a mean population of 1603 (standard deviation 394).

2.4.1. Exposure variations at different deprivation levels

We obtained socioeconomic and health data from the English Indices of Multiple Deprivation (IMD) 2015 (DCLG, 2015). The IMD dataset includes social deprivation scores for the LSOAs across seven sub-domains (i.e., income, health, employment, crime, education, living environment, and barriers to housing and services). In this study, we used IMD deciles to investigate variations of greenspace exposure in

LSOAs with differing deprivation levels. An IMD decile value of 1 indicated the highest level of deprivation and 10 indicated the lowest level of deprivation. We used a box plot to explore the differences in greenspace exposures visually and then conducted an ANOVA with post hoc Tukey Honestly Significant Difference analysis to identify whether the differences in exposure values among the IMD deciles were statistically significant or not.

2.4.2. Associations of greenspace exposures with YPLL

The years of potential life lost outcome variable was extracted from the IMD 2015 health and disability sub-domain. YPLL measures 'premature death,' defined as death before the age of 75 from any cause, including death due to disease or other external causes. This indicator was estimated based on mortality data for the period 2008 to 2012 as the numerator, and the denominator was the 2008–2012 population estimate. The details of YPLL estimation can be found in the IMD technical report (DCLG, 2015). YPLL is age and sex standardised, therefore it is not vulnerable to biases such as disproportionate representation by a certain demographic section, or double counting (DCLG, 2015; Dennis et al., 2020). Additionally, it is also weighted by the age of the individual who has died, reflecting the higher impact of an unexpected death of a younger person, than an older one. It should be noted that YPLL is a population-level inference of premature mortality, which has been aggregated to the neighbourhood level (LSOA) for the purposes of this analysis. The associations that we observe are therefore at the population-level and the effects are not transferable to individuals or to different spatial scales of analysis.

Pearson correlation coefficients (r) were calculated to assess correlations among greenspace exposure metrics, YPLL, and neighbourhood socioeconomic variables. Additionally, a hierarchical cluster-based 'heat map' visualisation was produced based on the correlation coefficients to identify clustering among correlated variables (method details in Zhang et al., 2017). Most previous studies have used non-spatial models to explain the associations between exposure and health outcomes (Mitchell and Popham, 2008; Dadvand et al., 2015; Dennis et al., 2020). In this study, we have applied both non-spatial and spatial models to investigate the associations between greenspace exposures and YPLL based on frequentist statistical inferences. Results from non-spatial regression models indicating associations between greenspace exposures and YPLL are presented in Note-2, *Supplementary document*. In our non-spatial models, we applied the generalised linear model (GLM) with a gamma distribution and an identity link function to account for positive values of YPLL that could skew distributions. We have also performed multiple linear regressions to aid comparison between the non-spatial and spatial models. While the GLM allows non-normal error distribution of the residuals, the residuals may have spatial autocorrelation, which might violate the assumption of independence of error distribution. We therefore computed Moran's I statistic (Moran, 1950) for the model residuals in order to identify the presence of spatial autocorrelation in non-spatial models. All the non-spatial models indicated significant spatial autocorrelation in the regression residuals (Note-2, *Supplementary document*), thus demonstrating spatial dependency and violating the model assumption of independence of observations and inflating the model parameter estimates. As a result, these results are not reported here.

We fitted spatially explicit models to account for potential spatial autocorrelations and spatial dependencies among the variables in the models. As such, we used a 'spatial autoregressive with autoregressive error' (SARAR) model, which accounts for spatial dependence in both independent and dependent variables (Kelejian and Prucha, 2010; Bivand et al., 2013). As both our dependent and independent variables were spatially explicit, SARAR provides a better fit for the model than other spatial regression models such as spatial lag and spatial error models. The details of SARAR modelling can be found in Note-3, *Supplementary document* and Kelejian and Prucha (2010). In our modelling we used a first order queen contiguity spatial weight matrix, this

decision was informed by a previous study see Anselin and Rey (2014) and our experimental observations of model fit.

We formulated several regression models sequentially for both the spatial approaches, in order to explore the associations between YPLL and greenspace exposure for all of the LSOAs (all population model). In this regard we considered unadjusted and adjusted models. Model 1 is un-adjusted where each greenspace exposure metric was entered in separate models; Model 2 was adjusted for confounders. Model 2 included variables from Model 1 and the confounders. The selection of confounders was informed by a previous systematic review of the relevant literature (Labib et al., 2020b) and reference to several original studies (Richardson and Mitchell, 2010; Wheeler et al., 2015; Dennis et al., 2020). In the fully adjusted models, we included income scores, crime deprivation scores, barriers to housing-service scores, distance to nearest general medical practice (GPs), the Shannon diversity score, and annual average PM₁₀ concentration as confounders. Income, crime, and barriers to housing and service scores acted as proxy indicators for income levels, crime rates, physical and financial accessibility of housing and local services together. Distance to GP represents access to health services (Ensor and Cooper, 2004; Daras et al., 2019). The Shannon diversity score indicated the overall land cover mix within the study area as a proxy for vegetation quality (Dennis et al., 2020). Finally, the PM₁₀ concentration indicated one element of the ambient air quality of these neighbourhoods since there may be an additional health burden from air quality which is not accounted for by greenspace benefits (Medina et al., 2004; Jeanjean et al., 2016; Khomenko et al., 2021). We also tested sensitivity to other confounders such as employment and the education deprivation score, both of these demonstrated high multicollinearity with income deprivation while not improving the models, and so were not included. We assessed each model's explanatory power (pseudo R²), Akaike Information Criterion (AIC), and likelihood ratio test (LR) in evaluating the model's goodness of fit, prior to determining the optimum selection of confounders and most robust models.

In addition to population analysis for all the neighbourhoods (LSOAs) in the study area, we explored whether the association between aggregate greenspace exposure and YPLL varied by sub-groups at different levels of deprivation. We used the IMD score to classify each LSOA, and hence its population, into multiple deprivation quintile groups (Table S2, *Supplementary document*). For each group, we applied the same modelling structure and adjusted for the same confounders as previously described for the pooled analysis. Statistical significance was set at $p < 0.05$, and the v.3.6 R (R Team, others and R Core Team, 2013) statistical package was used for all statistical analyses. The code and results can be found in GitHub (<https://github.com/labiblm/CompositeGreenspaceExposure>).

3. Results

3.1. Greenspace exposure distribution and variations

The greenspace exposure index (single and composite) maps are presented in Fig. 1. It is clear from Fig. 1 that greenspace availability (Fig. 1A) and visibility (Fig. 1C) have a similar broad spatial pattern: exhibiting high exposure values at the rural periphery of the study area and along riparian corridors; and low exposure values in the urban cores. However, while the broad patterns are similar, the detailed patterns differ considerably due to the presence of buildings or other obstructing features that affect greenspace visibility. Fig. 1 also illustrates that the accessibility index (Fig. 1B) exhibits a rather different spatial pattern to the other two measures. We found relatively higher accessibility exposure values in and around dense urban cores, likely because these areas typically have many small to medium-size public parks and gardens (e.g., sized 2–20 ha), and good transport links. However, these urban parks and gardens did not result in higher availability or visibility exposure values due to their fragmented pattern and

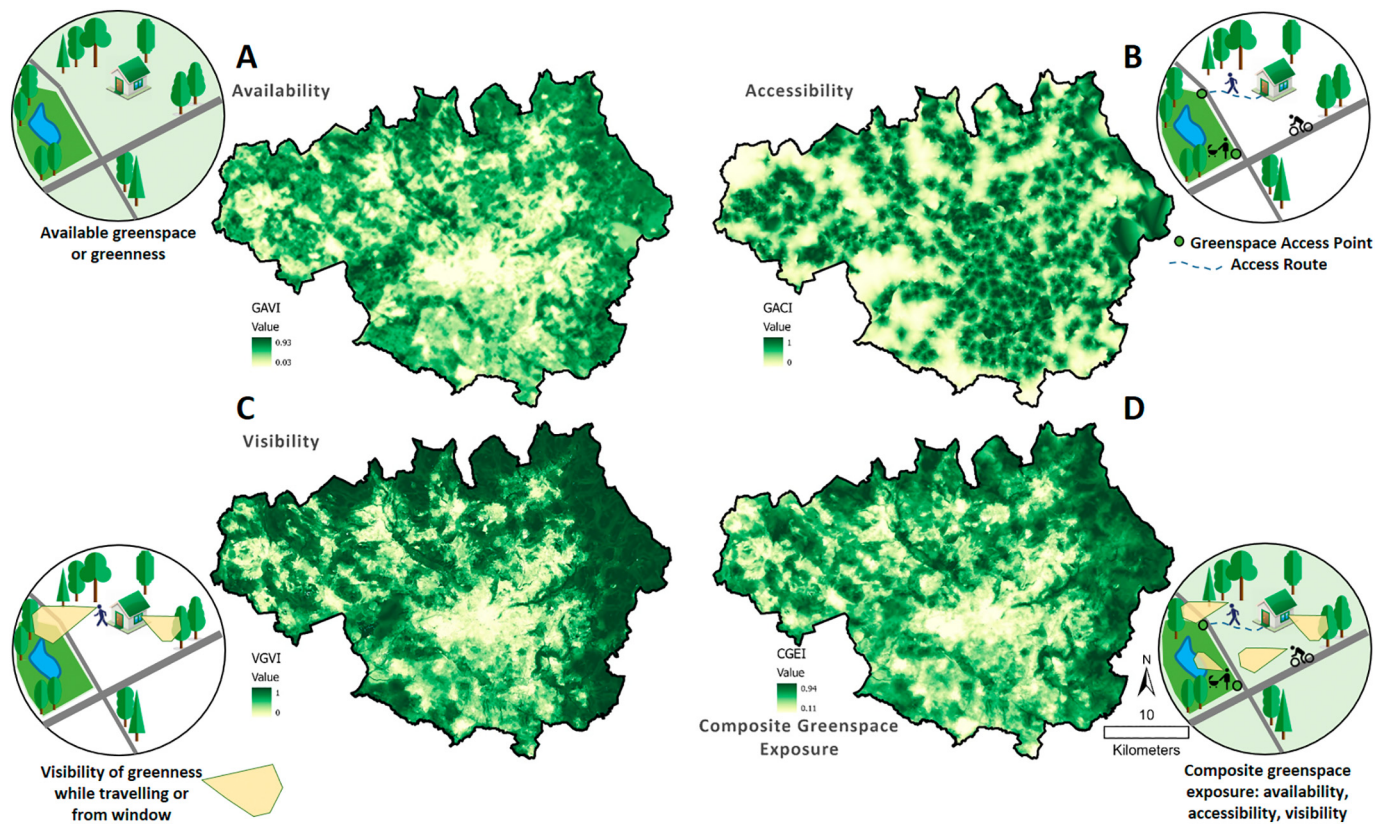


Fig. 1. Maps showing the single and composite greenspace exposure metrics (A) availability, (B) accessibility, and (C) visibility. (D) is the composite greenspace exposure map calculated by averaging three single exposure maps listed in (A–C). The conceptual representation of each type of exposure showed differences in measuring and modelling single exposure type and the composite is a summary of all three exposures together.

low connectivity. Fig. 1, therefore, demonstrates the differences in exposure values depending on how exposure is conceptualised and measured. Fig. 1(D) illustrates the composite greenspace exposure index map. This map is a combined output of the three single indices, and each cell of this map indicates the mean exposure value of the three exposure surfaces as our measure of aggregate greenspace exposure.

The relationship between the three input indices and the resulting composite index is also illustrated for a terraced house located near to a park (Fig. 2). Due to the close proximity of the park (only a few streets away), the house has a very high accessibility exposure value (0.957). Conversely it has a much lower visibility value (0.065) as the park cannot be seen due to the local terrain and intervening buildings. Finally, the house has a relatively low value for greenspace availability (0.225) because of the densely built-up nature of the surrounding area. The composite greenspace exposure index value ($CGEI = 0.416$) provides an aggregate overall moderate score for greenspace exposure for this particular house. Therefore, the composite greenspace exposure metric represents a comprehensive measurement that accounts for all three objective exposure types together.

3.2. Deprivation and greenspace exposures

The analysis of different greenspace exposure metrics and social deprivation deciles at the neighbourhood (LSOA) scale indicates variations in how greenspace exposures co-vary with social deprivation (Fig. 3). Fig. 3 illustrates that the least deprived areas have greater greenspace exposure compared to the most deprived areas in the domains of greenspace availability, visibility, and composite greenspace exposures. In the case of accessibility, however, we observed higher exposure values for more deprived areas, although with a smaller level of variation across the IMD deciles (summary descriptive statistics Table S3, Supplementary document). This result illustrates accessibility exposure

in our case study area has a different spatial distribution than availability and visibility based on existing greenspaces, parks and gardens. In the case of Greater Manchester, the most deprived neighbourhoods often have public parks and gardens in close proximity, but these neighbourhoods have low greenspace visibility and availability exposure because they are more densely built-up areas with low provision of private gardens and streets with low levels of canopy coverage. Conversely, the least deprived neighbourhoods usually exhibit lower building density, greater provision of private gardens, and streets with greater levels of canopy coverage, but with less provision of public parks and gardens in proximity. These results demonstrate that apparent inter-group variations in individual greenspace exposure measures can be misleading, and therefore illustrate the benefit of this composite measure.

The ANOVA analysis demonstrates that the differences in greenspace exposures among different deprivation deciles are statistically significant for all greenspace exposures (Table S4, Supplementary document). The post-hoc analysis identifies an interesting pattern in which differences are usually only significant in the case of the difference between the most and least deprived deciles. However, among similar deprivation deciles (e.g., 1–3 or 7–9) the differences in greenspace exposures are not statistically significant (Table S4, Supplementary document). This result implies that neighbourhoods that share similar deprivation levels usually demonstrate similar characteristics in terms of greenspace exposure.

3.3. Associations among greenspace exposure types and YPLL

Correlation analyses showed that there were very strong positive correlations among composite, availability and visibility metrics for the Greater Manchester neighbourhoods. By contrast, greenspace accessibility only had a small correlation with the composite metric, and a weak negative correlation with the availability and visibility exposure

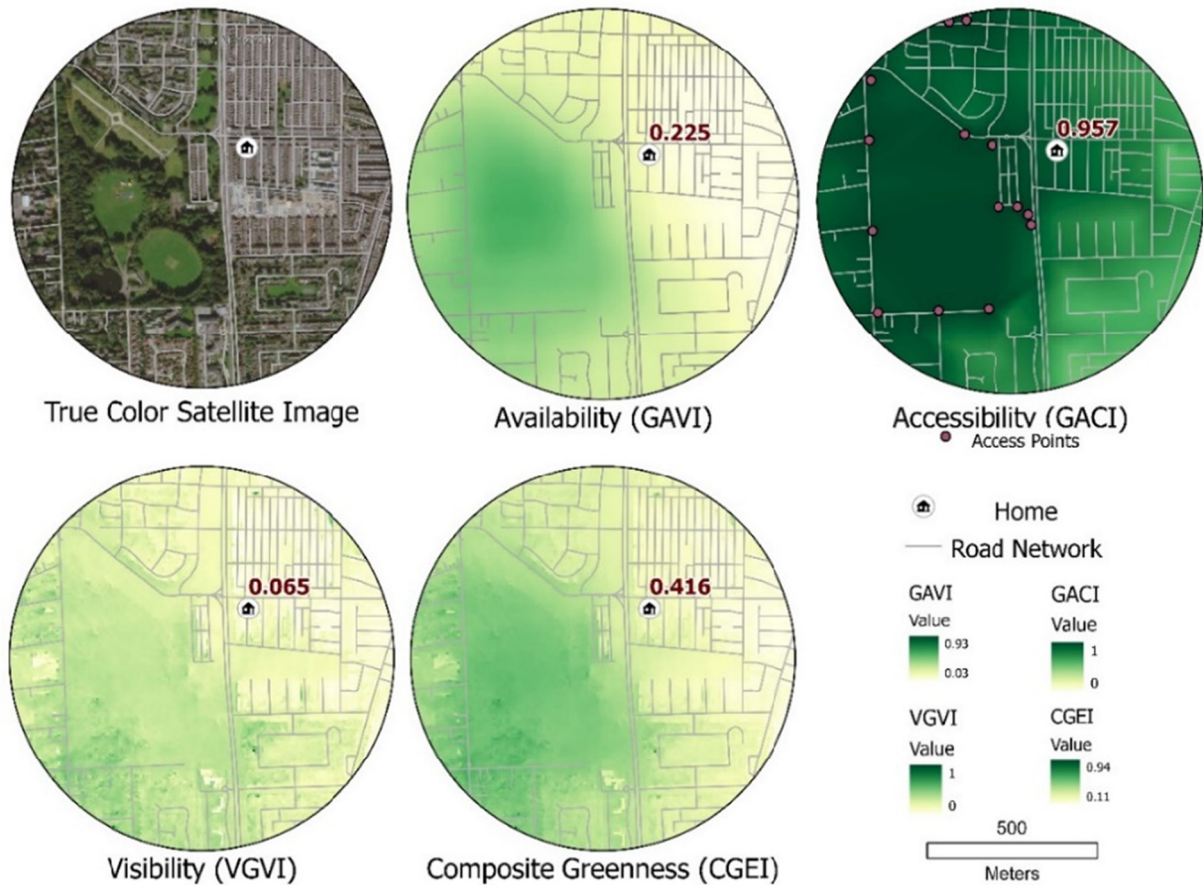


Fig. 2. Differences in greenspace exposure values for an example residence. GAVI shows availability exposure value, GACI shows accessibility exposure value, VGVI presents visibility exposure and CGEI shows the composite greenspace exposure value for the example residence.

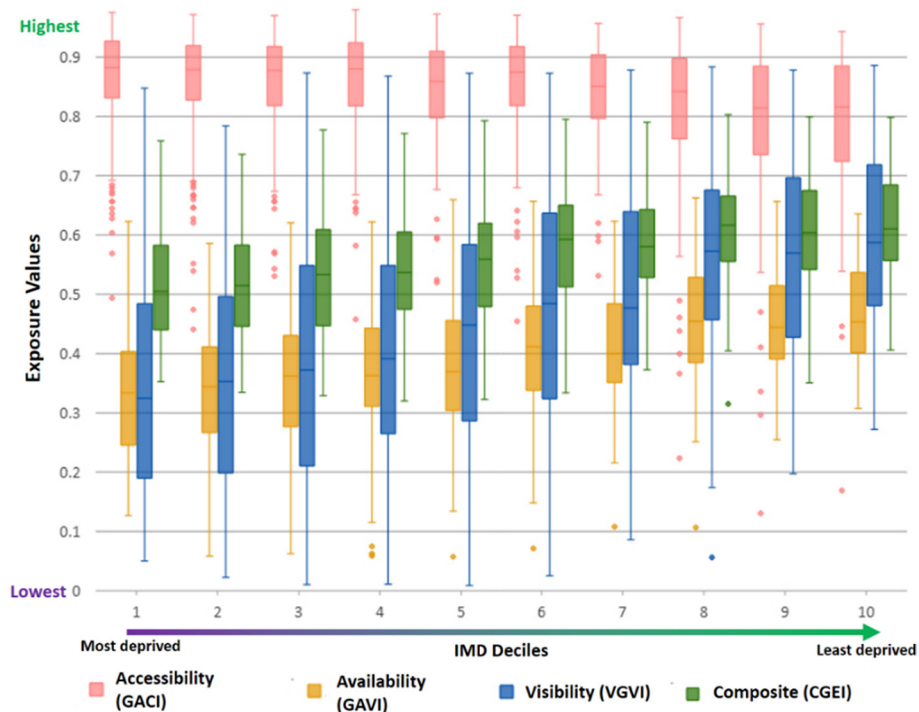


Fig. 3. Greenspace exposure indices compared to different IMD deciles. Dots indicate outliers in the measurements of greenspace exposures.

metrics (Fig. 4A). The ‘heat map’ also illustrates a similar pattern, as availability, visibility and the composite metric are clustered in the same group, indicating high similarity among these metrics; whereas the accessibility metric was not included in this cluster with the other greenspace metrics (Fig. 4B). Such divergent patterns of greenspace accessibility measure were also observed in the spatial distribution greenspace exposure maps shown in Fig. 1.

The correlation analyses and the correlation ‘heat maps’ further indicated that the availability, visibility, and composite greenspace exposure measures have moderate to weak negative correlations with YPLL and are in contrasting cluster groups with it as well (Fig. 4B). Additionally, these metrics have negative correlations with deprivation scores and the air quality variable, but they have positive correlations with distance to GP (Fig. 4A). In contrast, the accessibility greenspace exposure metric has a positive correlation with YPLL, deprivation score, and air quality variables, all of which are clustered in the same group (Fig. 4B).

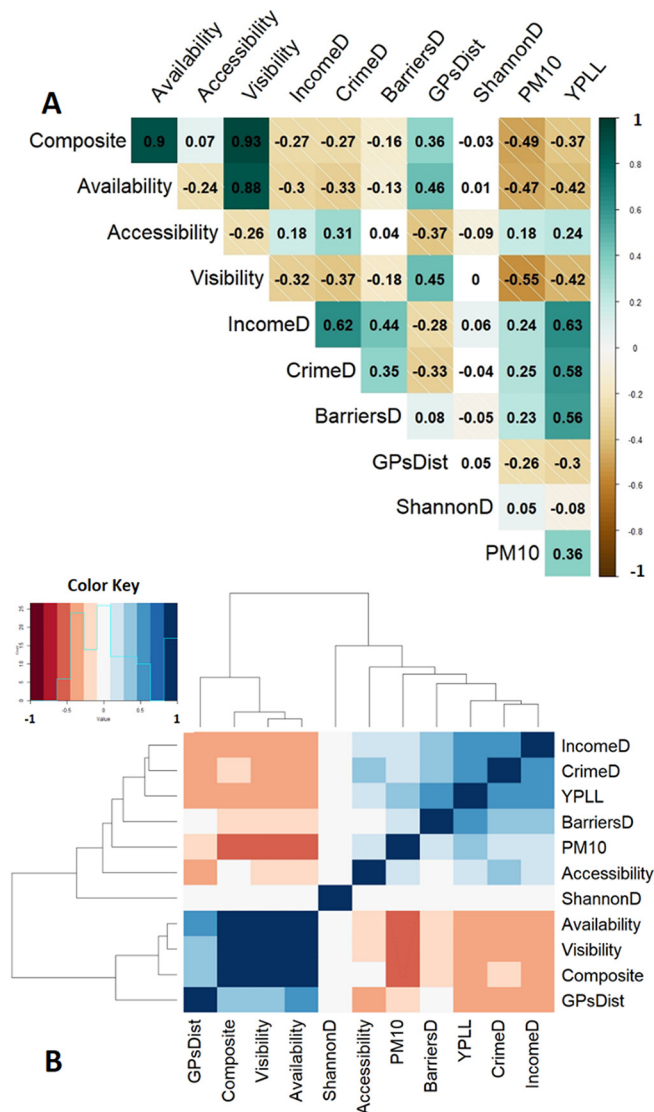


Fig. 4. (A) Correlations among greenspace exposure types, neighbourhood premature mortality (YPLL), socioeconomic variables, and the physical environment (i.e., land use mix, and air quality). Green boxes indicate positive correlation coefficients and brown boxes indicate negative correlation coefficients. Unshaded boxes indicate no significant correlations, shaded boxes indicate significance at $p < 0.05$. (B) correlation coefficient ‘heat map’ with clusters of variables. Blue boxes indicate similarity between clusters and observations, red boxes indicate dissimilarity of observations and dendrogram lines illustrate the hierarchical clusters among variables. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Regression analysis results are presented in Fig. 5 and Table 1. Similar to the correlation results, Fig. 5(A) indicates significant inverse associations of availability, visibility, and composite greenspace exposure with YPLL both for unadjusted and adjusted models. These associations are generally attenuated to some extent in the adjusted models. The results indicate increasing availability, visibility, and composite greenspace exposures were associated with lower YPLL values at the population level analysis. By contrast, for greenspace accessibility exposure, we found increasing accessibility was associated with higher YPLL values. In the unadjusted model, the effect is significant, but following the adjustment, the effect attenuates and becomes insignificant (Table 1). This result implies that, at the population level, accessibility exposures do not significantly relate to YPLL in the case study area. Overall, these findings meant that the association between greenspace exposure and YPLL, or premature mortality, differed across different types of greenspace exposure metrics. Therefore, measuring exposure using a single metric might not present a complete picture of the influence of greenspace exposure on health (e.g., premature mortality).

The adjusted CGEI exposure model indicates that a single unit increase (from CGEI 0 to 1, or 100%) in composite greenspace exposure is associated with a reduction of approximately ten years in the YPLL value. Thus, a 0.1 increment or 10% increase in the composite greenspace exposure value is associated with an approximately one year lower premature mortality value. It should be noted that both availability and the composite exposure models have nearly identical effect size and R^2 values (~0.702, Table 1), meaning that they both explain similar levels of variability in YPLL after controlling for confounders. This is partly a result of similarities between the availability and visibility metrics caused by the fact that ‘visible’ greenspace (a measure of line of sight) is also likely to be ‘available’, though the inverse is not necessarily true. The result of this is that the effect of the availability and visibility surfaces upon the composite are quite similar, meaning that (in the case of equal weighting of the three inputs in Eq. (4) the composite surface resembles both of them quite closely (Fig. 4A). In addition, the exposure metrics and health indicator were aggregated at neighbourhood scale. The aggregated information may have reduced variance, which in turn affected the explanatory power of the models. Based on these results, we therefore cannot fully confirm which metric is the better predictor of neighbourhood level YPLL for our study area. However, despite having very similar R^2 values, the composite metric when compared with all the individual exposure metrics (Fig. 5A), has a slightly greater effect size (Table 1). In addition, conceptually the CGEI measures holistic greenspace exposure. Considering the preceding we used the CGEI metric for subgroup analysis in the case of each socioeconomic deprivation group.

Fig. 5B shows variations in the effects of composite greenspace exposure on YPLL for people living in areas with different socioeconomic deprivation levels. Increasing greenspace exposure is significantly inversely correlated with premature mortality among people living in areas of greater deprivation (groups 3–4) compared to those in the least deprived areas (groups 1–2). It should be noted that there is no significant composite greenspace exposure effect on YPLL for people living in areas of relatively low or no deprivation (groups 2, 1). Additionally, for people living in the least deprived areas (group 1), the effect of composite greenspace exposure on YPLL is the opposite of the hypothesised direction, implying that in the least deprived areas, increased greenspace may be associated with increased YPLL, though this effect is not statistically significant. For group 5 greenspace exposure is not significant in the adjusted model. It is possible that in the most deprived areas, the higher premature mortality rate may be attributed to other reasons that we did not consider in the models (e.g., unhealthy food habits, race/ethnicity) (Cecchini et al., 2010). Nevertheless, overall we have observed a deprivation-related gradient in the effect of greenspace exposure on premature mortality.

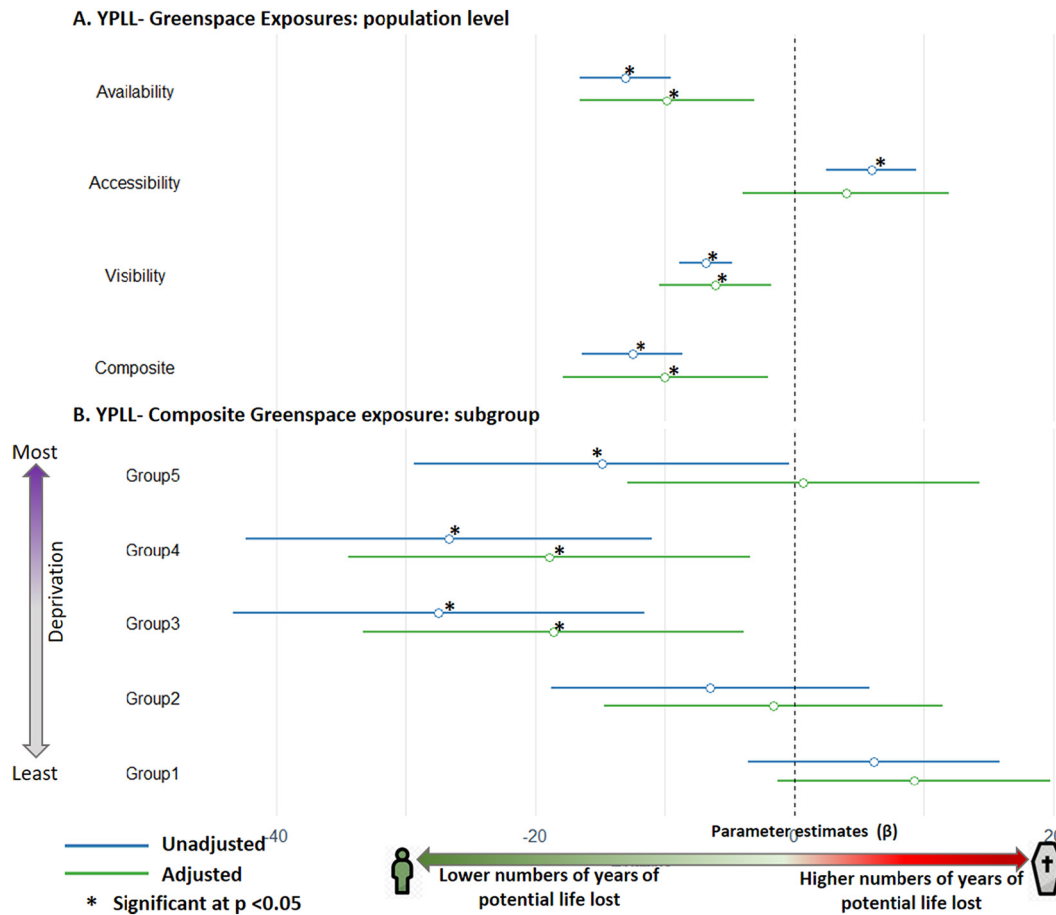


Fig. 5. The YPLL parameter estimates greenspace exposure relations for all neighbourhoods with a 95% confidence interval (A), and neighbourhoods stratified by different socioeconomic deprivation groups (B). The positive parameter estimate values indicate higher numbers of years of potential life lost, and negative values indicate lower numbers of years of potential life lost at the population and sub-group levels. (Detailed model outputs can be found in Note-3, *Supplementary document*).

4. Discussion

4.1. Multiple spatially explicit measures of greenspace exposures

To our knowledge, this is the first study that estimates greenspace exposure comprehensively by combining multiple spatially explicit greenspace exposure measures into a composite metric. Previously *Rugel et al. (2017)* developed a composite natural space index at a fixed spatial scale (i.e., postcode) combining multiple measures of greenness. The methodological approach that we have developed in this study has improved upon this fixed scale approach by standardising and harmonising different greenspace exposure measures in order to

produce a composite metric that estimates aggregate greenspace exposure at hierarchical spatial scales and fine spatial resolutions. We named this the composite greenspace exposure index (CGEI). We argue that the greenspace exposure modelling and mapping methods we developed in this study offer the opportunity to provide comprehensive estimates at the population level and provide the basis to infer individual-level aggregate greenspace exposures at meaningful temporal and spatial scales.

CGEI differs from traditional single exposure metrics in that it can capture objective greenspace exposure holistically and combine multiple inter-twined greenspace exposure types. Such a holistic exposure measure allows for a better understanding of the overall impact of

Table 1

Greenspace exposure and YPLL relationships using spatial regression models. The ‘Effect Estimate’ values are the numbers of years of potential life lost associated with unit greenspace exposure.

Greenspace exposure	Model adjustment	Model outcomes			R ²
		Effect estimate	95% CI (low; up)	p	
Availability	Unadjusted	-13.03***	-16.59; -9.48	6.661E-13	0.577
	Adjusted ^a	-9.8***	-15.90; -3.06	0.0037	0.702
Accessibility	Unadjusted	5.95***	2.47; 9.44	0.0008	0.565
	Adjusted	4.02	-3.94; 11.98	0.3228	0.70
Visibility	Unadjusted	-6.80***	-8.8; -4.80	2.64E-11	0.576
	Adjusted	-6.06***	-10.36; -1.75	0.0057	0.702
Composite	Unadjusted	-12.46***	-16.35; -8.56	3.621E-10	0.573
	Adjusted	-9.95***	-17.89; -2.00	0.014	0.702

CI: Confidence interval. R² is Nagelkerke pseudo-R-square. (Detailed model outputs can be found in Note-3, *Supplementary document*).

*** Indicate statistical significance at p < 0.05.

^a Model adjusted for: Income score, Crime score, Barriers to housing score, Shannon diversity score, Distance to nearest general medical practice, and annual average air pollution (PM₁₀).

greenspace exposure upon health outcomes such as mortality, morbidity, and non-communicable diseases that may be linked with multiple greenspace exposures. However, it should be noted that the composite is an aggregated expression of the single greenspace indices. Although it provides generalised exposure information it obviously cannot determine the specific cause-effect relations of each exposure type to specific health outcomes. For example, visibility may be more important in understanding the impact on a specific mental health outcome such as stress recovery, (Ulrich, 1984; Kaplan, 1995), than total greenspace exposure as represented in the composite measure. We therefore argue that the use of the CGEI does not diminish the importance of the use of individual exposure maps, as they provide the opportunity to understand specific types of exposure and, hence, which pathways correspond to which health indicators. As a result, these metrics are complementary to one another when evaluating the health impact of greenspace.

Greenspace exposure maps (both composite and single metrics) can be used to assess greenspace exposures for places where people live (e.g., home), work (e.g., office), or go about their daily activities (e.g., school, shops) (Fig. 6). In addition to these fixed locations, the results allow the estimation of individual exposure profiles by taking account of people's everyday movement (e.g., using GPS tracks) and the amount of time they spend in different locations (*details in Note-4, Supplementary document*). In Fig. 6, we provide the example of a hypothetical household of three people (two adults, one child) with a CGEI value of 0.60 for the home. However, each household member spends differing amounts of time in various locations, and each member has different travel routes. The child travels to school (CGEI 0.42) along streets exhibiting moderate levels of green cover. The child has an overall CGEI of 0.55 based on time spent at home, school, and during periods of movement. Similarly, the first adult (adult1) travels to the office (office-1) (CGEI 0.34), taking a route that exhibits relatively low greenspace exposures for an overall CGEI of 0.501. Finally, the second adult (adult2) travels to a different office (office-2) (CGEI 0.31) along a route with low greenspace exposure, providing an overall CGEI of 0.48.

In the illustrative example (Fig. 6), we demonstrate how the mapped outputs could be readily used to generate aggregate greenspace exposure estimates for individuals in households with different exposure profiles. In turn these individualised exposure profiles could be linked with individualised health data. While the composite provides an estimation of aggregate greenspace exposures, the single exposure maps can nevertheless provide crucial insights into the different types of greenspace exposure that an individual may experience at different spatiotemporal scales. This type of space-time exposure profiling has been utilised in other exposure studies relying on activity-based exposure measurements (Briggs, 2005; Zhang et al., 2018), but multiple exposure pathways have not been considered. Our new approach reduces existing methodological challenges in objective greenspace exposure assessments that rely on single exposure metrics, including measurement of exposure at fixed spatial scales, the uncertain geographic context problem, and issues with heterogeneous methods of exposure assessments across studies (Kwan, 2012; Turner et al., 2017; Helbich, 2018; Labib et al., 2020b).

4.2. Impact of greenspace exposure types on premature mortality and deprivation

We tested the effects of greenspace exposure in relation to neighbourhood-level socioeconomic deprivation in its relation to years of potential life lost (YPLL), a measure of premature mortality. We found that areas with higher availability and visibility exposures are usually the least deprived, and the most deprived areas have comparatively low greenspace exposures according to the same metrics. When we considered accessibility exposure. The results demonstrated the opposite trend, indicating that the most deprived areas have greater accessibility to public greenspaces (Dennis et al., 2020), than the least deprived due to the frequent proximity of small or medium size (e.g., 2–20 ha) public greenspaces. We speculate that such patterns could be attributed to the nature of the development of post-industrial towns in England, where parks were historically built by

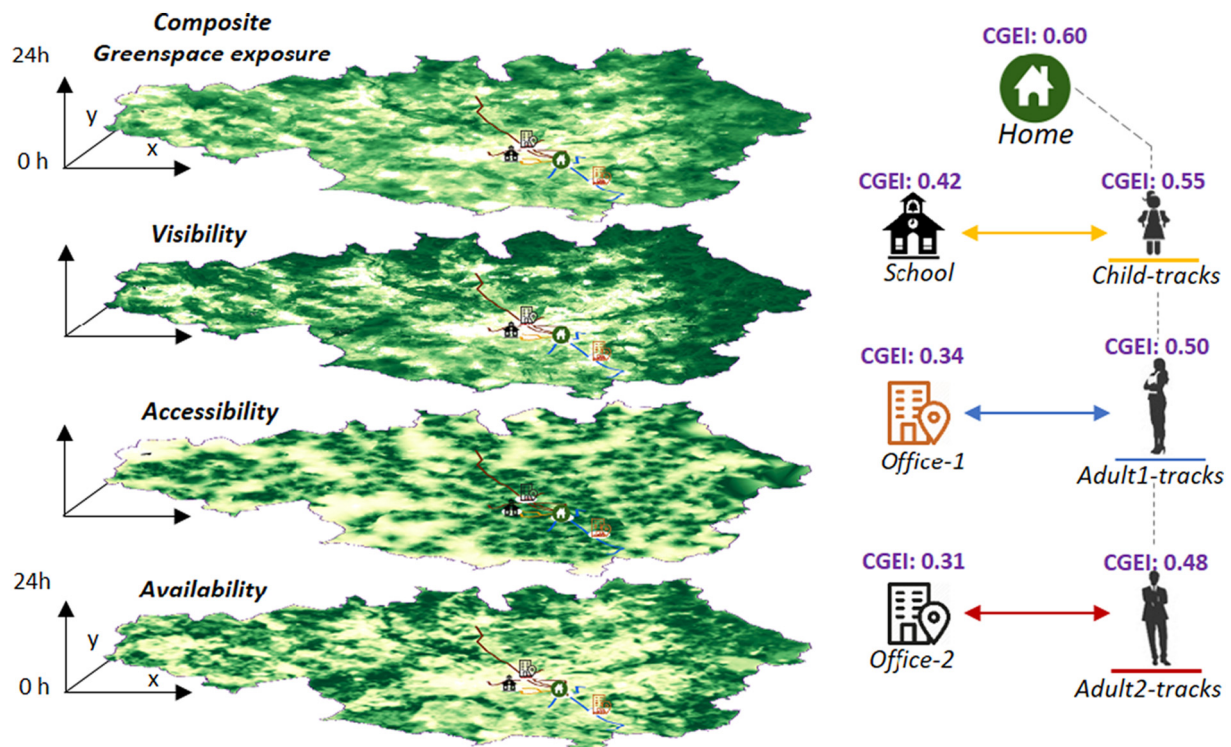


Fig. 6. A schematic illustration of space-time mobility-based greenspace exposures for three individuals living in the same household, and the varying exposure profiles each experience depending on the characteristics of their daily movement patterns.

mill owners among dense terraced housing in order to support the factory workers, following the Open Spaces Act of 1877 and Public Health Act 1875 (Clark, 1973; Jordan, 1994). In such towns and cities, such dense terraced housing areas typically present with higher deprivation values. Our accessibility result is therefore consistent with previous work suggesting that greenspace accessibility is greater in areas of greater socioeconomic deprivation in the UK (Barbosa et al., 2007; Jones et al., 2009).

When comparing the mean composite greenspace exposure values for neighbourhoods at differing socioeconomic levels, we observed a deprivation-related gradient in greenspace exposure, in which the least deprived areas exhibited higher composite greenspace exposures than the most deprived areas. These results align with previous research showing that greener neighbourhoods are also less deprived communities (Mitchell and Popham, 2008; Rigolon et al., 2021). However, we show that the composite measure presented here can estimate and differentiate greenspace exposures to a greater extent than the individual single exposure metrics which are typically used to evaluate the relationships between inequality and greenspace exposure. The results of the present study also suggest that the composite index has the potential to reveal new or refined health associations in previously reported studies which have so far only used traditional (e.g., percentage of neighbourhood greenspace) or individual metrics (e.g., accessibility). As we have demonstrated, single exposure metrics measuring differing factors such as accessibility or availability may provide results that appear different or contradictory (Fig. 3). As a result, single metric-based results can potentially misrepresent the true level of differences in greenspace exposure experienced by neighbourhoods at different levels of socioeconomic deprivation.

When we evaluated the associations between greenspace exposure and premature mortality (i.e., YPLL), the results showed that individual greenspace exposure metrics (i.e., availability, accessibility, and visibility) provided considerable variances in effect sizes (e.g., effect size for availability: -9.80 , visibility: -6.06) and directions regarding the associations between exposures and outcomes (Table 1). Both the availability and visibility greenspace exposure metrics indicate significant negative associations and imply that higher levels of greenspace availability or visibility are associated with lower YPLL (reduced premature mortality). These findings are consistent with previous studies indicating greater availability or visibility of greenspace can have positive health benefits, including reductions in mortality (Fong et al., 2018; Kondo et al., 2018; Rojas-Rueda et al., 2019).

Our results, which show no significant associations between greenspace accessibility and YPLL, are consistent with several previous studies reporting associations between greenspace exposure and health outcomes (Richardson et al., 2010; Klompaker et al., 2018; Jarvis et al., 2020; Labib et al., 2020) but not with others (Coutts et al., 2010; Wilker et al., 2014; Dennis et al., 2020). These latter studies reported positive health effects related to increased greenspace accessibility. Such inconsistent results might be linked to different approaches to conceptualising and measuring accessibility (e.g., shortest distance vs. fixed distance) between the present study and previous studies (Jarvis et al., 2020; Labib et al., 2020b). It must also be noted that several other studies have argued that the positive effect of greenspace accessibility may be context-specific (Jones et al., 2009; Richardson et al., 2010). For example, proximity to greenspaces does not always ensure greater use in many communities if areas have poor walkability (Zuniga-Teran et al., 2019) or where the poor perception of greenspace may deter its use (Jones et al., 2009). Therefore, even if greenspace accessibility is high, if that greenspace attracts few visits then the associated proactive health benefits might not be significant.

Despite these inconsistent associations when using single metrics, in line with our hypothesis, we found increasing overall community greenspace (i.e., measured as composite greenspace exposure) was associated with a lower level of premature mortality. Our overall findings are consistent with previous reports of the positive impact of greenspace

exposure associated with reductions in different types of mortality among people living in greener communities (Mitchell and Popham, 2008; Coutts and Horner, 2016; James et al., 2016; Rojas-Rueda et al., 2019).

Our research suggested that variations in the effects of different greenspace exposure measurements might link with the underlying mechanistic pathways associated with each exposure metric (Marselle et al., 2021; Zhang et al., 2021). For example, greenspace availability (i.e., the overall amount of vegetation in an area) might reduce the effect of harmful environmental stressors such as air pollution or heat (Markevych et al., 2017; Lindley et al., 2019) and reduce premature death. Greenspace accessibility (i.e., the relative proximity of public green spaces) may be associated with building capacities by encouraging physical activity and social cohesion (Hartig et al., 2014; Markevych et al., 2017). However, in our case study area at the population level, such mechanisms might not be as effective as other pathways. This conclusion was supported by the insignificant association between accessibility and YPLL in the adjusted model. Visibility of greenspace may be linked with restoring capacities related to psychological mechanisms such as stress recovery that reduce morbidity (Ulrich, 1984; Frumkin et al., 2017; Markevych et al., 2017). In the present study, the effect size of visibility was even smaller than that for availability, indicating the positive effect of visible greenspace might have an even lower relative impact than availability in lowering premature mortality. Overall, the composite greenspace exposure showed a slightly larger effect size than any individual metric in the adjusted model (Table 1). As the composite greenspace exposure accounts for all three well-known important mechanistic pathways through which nature influences health (Markevych et al., 2017), it is reasonable to argue that composite greenspace exposure might have a stronger relationship with premature mortality than has previously been considered when based only on individual metrics (e.g., visibility, accessibility). This finding is critical because it implies that many health-related indicators such as mortality, morbidity, and other non-communicable diseases and health syndromes might be influenced by the aggregate impact of multiple greenspace exposure types due to the multiple associated exposure benefit pathways.

We also found significant differences in the relationship between composite greenspace exposure and YPLL among groups at different socioeconomic deprivation levels. Our results indicate that increasing greenspace exposure in areas of moderate to extreme deprivation corresponds to a lower YPLL (reduced premature death) than in the least deprived areas. These findings are consistent with published studies that suggest that the health benefits of increased greenspace exposure are more pronounced among low-income and socioeconomically deprived populations (Marmot, 2020; Rigolon et al., 2021; Wu and Kim, 2021).

4.3. Methodological approach in multiple spatially explicit exposure modelling

In this study, we developed a new methodological solution for measuring aggregate greenspace exposures at hierarchical scales and fine resolutions. The complete measurement of a set of environmental exposures can be linked with the broader framework of the 'Exposome,' which considers the totality of human environmental exposures from conception onward (Wild, 2012). The exposome concept includes all external and internal exposures throughout a human lifespan (Wild, 2012; Vrijheid et al., 2020). Analysis of a subset of external greenspace exposures might provide a more useful framework for studying the associations between aggregate greenspace exposure and human health, as opposed to using a single type of greenspace exposure. The methodological process we introduce in this study illustrates an approach that allows a measurement of the summation of the multiple spatially explicit exposures and allows capturing high spatial-temporal exposure variability at multiple scales. Therefore, a spatially explicit composite measurement could be considered a step toward a "spatial exposome" and results from such measurements can be utilised to support the

broader exposome framework. A spatial exposome allows for spatial hierarchies in aggregate exposure assessments for any given set of external environmental exposures, which can then, in turn, allow the inference of comprehensive exposure profiles for individuals as well as for populations of varying sizes. While our composite exposure metric provides an introductory example of the spatial exposome concept, it will require further development in terms of other sets of external exposure variables (e.g., air, noise pollution) and time points.

4.4. Limitations and future developments

Although we have evaluated the robustness of our methodology and elucidated the implications for understanding premature mortality and socioeconomic inequality, our study does have some limitations. First, our exposure metrics are spatially explicit objective measurements of greenspace exposure based on the different vegetated areas alone (i.e., they do not include blue spaces). Furthermore, the measurements primarily represent the quantity of greenspace rather than its perceived quality (which we have used as a confounder). As a result, consideration should be given to integrating objective (e.g., vegetation diversity) or subjective (e.g., attractiveness) measurements of quality with the composite greenspace exposure index in future studies. Such subjective perceptions of greenspace exposure could be collected using individual surveys and integrated with our objective assessment to provide an overall understanding of the effect of both objective and perceived greenspace exposure on health.

Second, to evaluate our new metric, we investigated the direct association between greenspace exposure and premature mortality. We did not conduct an explicit mediation analysis to understand the mediating effect of variables such as air or noise pollution or heat. Although some previous studies have indicated mixed results for, and relatively lower effects from, these variables in terms of their mediation of the associations between greenspace exposure and health (James et al., 2016; Vienneau et al., 2017). In future studies, these variables should be critically evaluated for their potential effects on results generated using our composite greenspace measurement approach applying more explicit mediation and moderation approaches, such as structural equation modelling (Dzhambov et al., 2020).

Third, we used equal weights for each of the single greenspace exposure metrics in the composite greenspace calculations. No pre-existing evidence existed to indicate the precise level of health benefits associated with each type of greenspace exposure based on their differing effect pathways, the individualistic nature of the impacts, and the spatiotemporal variability. However, it would be possible to construct a weighted average for the composite exposure estimation if such information is available. In *supplementary Note-1*, we provided a worked example of weighted composite greenspace exposure index estimation based on a range of hypothetical weights. To generate actual weights for different input layers of the composite in future applications, a meta-analysis or analytical hierarchy approach could be adopted to identify the relative importance of availability, accessibility, and visibility exposures considering different health outcomes.

Finally, our population-level inference of greenspace exposure and premature mortality was made using cross-sectional data, and we had no means of knowing the extent to which individuals experienced variances in their levels of greenspace exposure throughout their lives. Therefore, we cannot confirm that the relations we observed can be inferred as a causal effect of greenspace exposure on premature mortality. Additionally, the population and exposure information were aggregated at the neighbourhood administrative boundary scale (LSOA). As a result, the correlations observed in aggregate cannot be transferred to individuals or to different spatial scales of analysis, due to the potential presence of zoning effects associated with the modifiable areal unit problem and ecological fallacy (Openshaw, 1984; Goodchild, 2011). Although Annerstedt Van Den Bosch et al. (2016) noted that spatially aggregated population data is an acceptable alternative to using individual

data, we recognise that to fully understand the effect of comprehensive greenspace exposure on individual health, longitudinal and life-course assessments of greenspace exposure need to be accounted for (Turner et al., 2017; Jia, 2019).

5. Conclusion

In this study, we provided a new methodology for estimating multiple greenspace exposure metrics. Our approach has the potential to reduce the heterogeneity inherent in current approaches to greenspace measurements, as it provides a technique that allows multiple quantitative greenspace exposure measurement practices to be harmonised and standardised. This enables a shift from the traditional approach of measuring single greenspace exposure toward a more holistic approach to greenspace exposure estimation at multiple spatial scales. Our approach also allows more spatially explicit aggregate exposure assessment and can be partially linked with the exposome framework. When applying the new index to population-level analyses, we found significant negative associations between community greenspace exposure and premature mortality. The association between increased greenspace exposure and lower premature mortality was particularly marked in areas with higher aggregate greenspace exposure levels. However, such associations differed significantly based on any given community's level of socioeconomic inequality. Overall, our results suggest that composite greenspace exposure measurements can reveal new insights into the public health benefits of exposure to the natural environment. Furthermore, the variations in the level of benefit derived by individuals from given levels of greenspace exposure may depend on the population of which they are members and their socioeconomic status.

CRediT authorship contribution statement

S.M. Labib: Conceptualization, Methodology, Data curation, Formal analysis, Investigation, Software, Writing – original draft, Visualization, Writing – review & editing. **Sarah Lindley:** Conceptualization, Methodology, Supervision, Resources, Methodology, Writing – review & editing, Funding acquisition. **Jonny J. Huck:** Conceptualization, Methodology, Software, Resources, Supervision, Writing – review & editing.

Declaration of competing interest

We declare that we have no conflicts of interest.

Acknowledgement

We would like to thank the reviewers and editor of this paper for their constructive comments and suggestions. S.M. Labib is funded by the School of Environment, Education and Development Postgraduate Research Scholarship at the University of Manchester. Elements of this work were carried out as part of the Green Infrastructure and the Health and Well-Being Influences on an Ageing Population (GHIA) project (2016–2019) <http://www.ghia.org.uk>. This study was funded by the Natural Environment Research Council, the Arts and Humanities Research Council and the Economic and Social Research Council under the Valuing Nature Programme. NERC grant reference number NE/N013530/1.

Appendix A. Supplementary Document

Supplementary document to this article can be found online at <https://doi.org/10.1016/j.scitotenv.2021.147919>.

References

- Annerstedt Van Den Bosch, M. et al. (2016) 'Development of an urban green space indicator and the public health rationale', *Scandinavian Journal of Public Health*. SAGE Publications/Sage UK: London, England, vol. 44(2), pp. 159–167. <https://doi.org/10.1177/1403494815615444>.

- R Team, C., others and R Core Team (2013) 'R: A Language and Environment for Statistical Computing'. Available at: <https://www.gbif.org/tool/81287/r-a-language-and-environment-for-statistical-computing> (Accessed: 10 September 2020).
- Richardson, E. et al. (2010) 'The association between green space and cause-specific mortality in urban New Zealand: an ecological analysis of green space utility', *BMC Public Health*. <https://doi.org/10.1186/1471-2458-10-240>.
- Remme, R.P., Frumkin, H., Guerry, A., King, A., Mandlea, L., et al., 2021. An ecosystem service perspective on urban nature, physical activity, and health. *Proc. Natl. Acad. Sci. U. S. A.* 118 (22), e2018472118. <https://doi.org/10.1073/pnas.2018472118>.
- Richardson, E.A., Mitchell, R., 2010. Gender differences in relationships between urban green space and health in the United Kingdom. *Soc. Sci. Med.* 71 (3), 568–575. <https://doi.org/10.1016/j.socscimed.2010.04.015>.
- Rigolon, A., et al., 2021. Green space and health equity: a systematic review on the potential of green space to reduce health disparities. *J. Environ. Res. Public Health* 18, 2563. <https://doi.org/10.3390/ijerph18052563>.
- Rojas-Rueda, D., et al., 2019. Green spaces and mortality: a systematic review and meta-analysis of cohort studies. *The Lancet Planetary Health*. World Health Organization 3 (11), e469–e477. [https://doi.org/10.1016/S2542-5196\(19\)30215-3](https://doi.org/10.1016/S2542-5196(19)30215-3).
- Rugel, E.J., et al., 2017. Beyond the normalized difference vegetation index (NDVI): developing a natural space index for population-level health research. *Environmental Research*. Academic Press Inc. 159, 474–483. <https://doi.org/10.1016/j.envres.2017.08.033>.
- Silva, R.A., Rogers, K., Buckley, T.J., 2018. Advancing environmental epidemiology to assess the beneficial influence of the natural environment on human health and well-being. *Environmental Science and Technology*. American Chemical Society 52 (17), 9545–9555. <https://doi.org/10.1021/acs.est.8b01781>.
- Stas, M., et al., 2021. Exposure to green space and pollen allergy symptom severity: a case-crossover study in Belgium. *Sci. Total Environ.* 781, 146682. <https://doi.org/10.1016/j.scitotenv.2021.146682>.
- Turner, M.C., et al., 2017. Assessing the exposome with external measures: commentary on the state of the science and research recommendations. *Annu. Rev. Public Health* <https://doi.org/10.1146/annurev-publhealth-082516-012802>.
- Ulrich, R. S. (1984) 'View through a window may influence recovery from surgery', *Science*. <https://doi.org/10.1126/science.6143402>.
- Vienneau, D., et al., 2017. More than clean air and tranquility: residential green is independently associated with decreasing mortality. *Environ. Int.* <https://doi.org/10.1016/j.envint.2017.08.012>.
- Vrijheid, M., et al., 2020. Early-life environmental exposures and childhood obesity: an exposome-wide approach. *Environmental Health Perspectives* <https://doi.org/10.1289/EHP5975>.
- Wheeler, B.W., et al., 2015. Beyond greenspace: an ecological study of population general health and indicators of natural environment type and quality. *Int. J. Health Geogr.* <https://doi.org/10.1186/s12942-015-0009-5>.
- Wild, C.P., 2012. The exposome: from concept to utility. *Int. J. Epidemiol.* <https://doi.org/10.1093/ije/dyr236>.
- Wilker, E.H., et al., 2014. Green space and mortality following ischemic stroke. *Environ. Res.* <https://doi.org/10.1016/j.envres.2014.05.005>.
- Wu, L., Kim, S.K., 2021. Health outcomes of urban green space in China: evidence from Beijing. *Sustain. Cities Soc.* 65. <https://doi.org/10.1016/j.scs.2020.102604>.
- Zhang, J., et al., 2021. Deciphering the changes in residential exposure to green spaces: the case of a rapidly urbanizing metropolitan region. *Build. Environ.* 188, 107508. <https://doi.org/10.1016/j.buildenv.2020.107508>.
- Zhang, L., et al., 2018. Impacts of individual daily greenspace exposure on health based on individual activity space and structural equation modeling. *Int. J. Environ. Res. Public Health* <https://doi.org/10.3390/ijerph15102323>.
- Zhang, Z., et al., 2017. Hierarchical cluster analysis in clinical research with heterogeneous study population: highlighting its visualization with R. *Annals of Translational Medicine*. AME Publishing Company 5 (4), 9. <https://doi.org/10.21037/atm.2017.02.05>.
- Zuniga-Teran, A.A., et al., 2019. Exploring the influence of neighborhood walkability on the frequency of use of greenspace. *Landsch. Urban Plan.* <https://doi.org/10.1016/j.landurbplan.2019.103609>.